G- Journal of Environmental Science and Technology

(An International Peer Reviewed Research Journal)

Available online at http://www.gjestenv.com

Immunoinformatics Approach to T-cell Epitopes in Human Immunodeficiency Virus

Deepa Chaturvedi, Mallika Pandey, Vijai Singh and Ravi Deval*

Department of Biotechnology, Invertis University, NH-24, Bareilly – 243123 U.P INDIA

ARTICLE INFO

Received: **02 Dec 2013**
Revised : **12 Dec 2013**
Accepted: **21 Dec 2013**

***Key words*:**
HIV, AIDS, Virulence Protein, Epitope, MHC, Vaccine

ABSTRACT

Human immunodeficiency virus (HIV) is major human disease which belongs to lentivirus (retrovirus). It is a major cause of acquired immunodeficiency syndrome (AIDS), a condition in humans in which progressive failure of the immune system allows life threatening opportunistic infections and cancers to thrive. A number of ways for transmission of HIV form human to another human through the blood, semen, vaginal fluid, pre-ejaculate or breast milk has been proved. An urgent need arises to establish antigen based immunodiagnostic for earlier monitoring of HIV and development of vaccines. In the present study, we have identified two major proteins i.e. ENV and NEF for identification of T-cell epitopes. We used two well established immunoinformatics tools Propred and Propred1 for identification. We identified a novel T-cell epitope for major histocompatibility complex class I and II with highest binding affinity. These finding provide a new insight for development of antigen based diagnostic kit and peptide based vaccine designing for controlling of AIDS.

1) INTRODUCTION

Human immunodeficiency virus (HIV) is a single stranded RNA virus which is responsible to cause of acquired immune deficiency syndrome (AIDS). AIDS is defined in terms of either a $CD4^+$ T cell count below 200 cells per microlitre or the occurrence of specific diseases in association with an HIV infection. The immunity of human body is weak. [1]. In a recent report suggested by Centers for Disease Control and Prevention (CDC) show that 1,148,200 Americans aged 13 and older were living with HIV [2]. However, in the 2008-09 National AIDS Control Organization has reported that 2.39 million people live with HIV/AIDS in India [3]. While a more recent investigation by the Million Death Study Collaborators in the British Medical Journal and they estimated that 1.4-1.6 million people suffer with AIDS [4]. There is an urgent need for controlling of AIDS by developing rapid, eco-friendly and cost effective immunodiagnostic capability. Therefore, two types of HIV have been characterized including HIV-1 and HIV-2. HIV-1 is the virus that was initially discovered and considered to be a more virulent, which is caused the majority of HIV infections worldwide [5]. The lower infectivity of HIV-2 compared to HIV-1 implies that fewer of those exposed to HIV-2 will be infected per exposure, because of its relatively poor capacity for transmission; HIV-2 is largely confined to West Africa [6].

The culture of HIV is in laboratory is difficult and time consuming process; thus, a number of immunoinformatics tool have been developed for prediction of T and B cells epitopes any major virulent proteins. These tools have been developed on the basis of accessible and validated data which having specific algorithms [7]. Identification of T-cell epitopes in proteins of Human immunodeficiency virus (HIV) by using Propred [8] and Propred1 [9]. We had also used the Propred and Propred1 tools because these tools had covered maximum number of alleles. This is one of the very recent technique which is used for the identification of antigenic peptide and can be further used in diagnostic and vaccines development. The aim of present study was to identify and map the T-cell epitopes in the major putative proteins of HIV type 1.

2) MATHODOLOGY

The sequence of two proteins i.e. NEF and ENV, of HIV type 1 were retrieved from NCBI-GenBank and expected molecular weight and isoelectric point (pI) value of both the proteins were verified using Gene Runner and ExPasy software. Propred [8] and Propred1 [9] immunoinformatics tools were used for the identification of the T-cell epitopes in the protein sequence of NEF and ENV. These tools cover maximum number of human leukocyte antigen (HLA) comparison as compared to the other immunoinformatics tools. During epitopes prediction, we had considered 4% threshold with selecting 30 as the maximum binding score to HLA molecules. Finally, the predicted T-cell epitopes were used for

* Corresponding Author: Dr. Ravi Deval
Email address : ravi.d@invertis.org

confirmation and homology matching with other amino acids by NCBI protein blast for the proportional study.

3) RESULT AND DISCISSION

HIV is global health problem and suppressed the human immune system that allows more susceptibility of infection. In this study, two proteins have a locus such as HIV-1 GP-1 was used for physicochemical analysis such as molecular weight, isoelectric point (Pi value) and antigenic nature. ENV protein showed the highest molecular weight 97.03kDa. Isoelectric point of these proteins was ranged between 5.99-8.88 .The physicochemical properties of the two proteins of HIV type-1 were shown in **Table 1**.

In the present study, we have used Propred tool for MHC class I prediction of T-cell epitopes in the proteins which include ENV and NEF, encoded by genome of HIV. The predicted epitopes, their amino acid position in the protein and their number of MHC class I binding allele are shown in **Table 2**.
Additionally, we have used Propred-1 tool for MHC class II prediction of T-cell epitopes again in ENV and NEF proteins. The predicted epitopes, their amino acid position in the protein and their number of MHC class II binding allele are shown in **Table 3**.
Similar approach has been previously used for T-cell epitopes prediction from human papillomavirus [10], tick borne encephalitis virus [11], influenza A virus [12] and mycobacterium tuberculosis [13].

Table 1: Physicochemical properties of different putative proteins of HIV TYPE 1

Protein Designation	Accession Number	Molecular Weight (KDa)	PI
ENV	NP_057856	97.23	9.17
NEF	NP_057857	23.47	5.99

Table 2: The predicted epitopes in different putative proteins of HIV type 1 with reference to MHC molecules class I.

Protein	T-cell epitope	Amino acid position	Number of MHC Class I binding allele
ENV	VTENFNMWK	89-97	1
	KEYAFFYKL	171-179	2
	SFEPIPIHY	209-217	1
	KLTSCNTSV	192-200	2
	KLTPLCVSL	121-129	1
	QMHEDIISL	103-111	1
NEF	AVDLSHFLK	84-92	1

Protein	T-cell Epitope	Amino acid position	Number of MHC Class II binding allele
ENV	LRIVFAVLS	694-702	49
	VQLNTSVEI	285-293	13
	VVKIEPLGV	487-495	27
	FIMIVGGLV	684-692	7
	IRLVNGSLA	745-753	6
	IKQLQARIL	572-580	4
	IRIQRGPGR	306-314	8
	IRPVVSTQL	250-258	12
	LLNATAIAV	813-821	5
	VSLLNATAI	811-819	2
	VGGLVGLRI	688-696	4
	VVQGACRAI	831-839	3
	MRQAHCNIS	325-333	19
	LVGLRIVFA	691-699	25
	LGFLGAAGS	519-527	19
	LRAIEAQQH	555-563	13
	WKNDMVEQM	95-103	11
	IVGGLVGLR	687-695	3
	VRQGYSPLS	707-715	4
	LKNDTNTNS	133-141	14
	YKLDIIPID	176-184	4
	VTIGKIGNM	317-325	2
	LVNGSLALI	747-755	9
	MLLGMLMIC	19-27	9
	VVLVNVTEN	83-91	13
	VGIGALFLG	512-521	1

	LIHSLIEES	640-648	6
	LLIVTRIVE	774-782	13
	LFIMIVGGL	683-691	25
	LKNSAVSLL	806-814	3
	IRHIPRRIR	839-847	17
	IVFAVLSIV	696-704	3
	LRSLCLFSY	759-767	14
	VVIRSVNFT	269-277	26
	VLVNVTENF	84-92	2
	IRCSSNITG	442-450	5
	IHYCAPAGF	214-222	4
	LWYIKLFIM	678-686	1
	VQGACRAIR	832-840	4
	IVGGLVGLR	687-695	3
	FVTIGKIGN	316-324	16
	LLSGIVQQQ	543-551	6
	LSIVNRVRQ	701-709	9
	FNISTSIRG	158-167	10
	YLKDQQLLG	585-593	2
	FGNNKTIIF	352-360	8
	YIKLFIMIV	680-688	11
	YKLTSCNTS	190-198	7
	WNASWSNKS	609-617	3
	IKQIINMWQ	419-427	13
	IGKIGNMRQ	319-327	11
	IVQQQNNLL	547-555	9
	FFYCNSTQLF	382-390	3
	FAILKCNNK	222-230	6
	LKYWWNLLQ	792-800	17
	LLLIVTRIV	773-781	19
	VEINCTRPN	291-299	5
	VLSIVNRVR	700-708	3
	IWNHTTWME	621-629	5
	VGLRIVFAV	692-700	16
	INNYTSLIH	634-642	11
	LGIWGCSGK	592-600	1
	IVQLNTSVE	284-292	10
	LGMLMICSA	21-29	19
	MLMICSATE	23-31	9
	LMICSATEK	24-32	8
	LLGMLMICS	20-28	7
	YYGVPVWKE	38-46	4
	LLLNGSLAE	258-265	7
	WFNITNWLW	671-679	1
	VIRSVNFTD	270-278	1
	VFAVLSIVN	697-705	1
	IIVQLNTSV	283-291	2
	VQQQNNLLR	548-556	1
	FNSTWSTEG	395-403	1
	FNITNWLWY	672-680	3
	MICSATEKL	25-33	3
	FQTHLPTPR	716-724	3
	LIVTRIVEL	775-783	2
	YKVVKIEPL	485-493	2
	WYIKLFIMI	679-687	5
	YAFFYKLDI	172-180	20
	FYKLDIIPI	175-183	7
	VNVTENFNM	86-94	3
	WLWYIKLFI	677-685	23

	VNGSLALIW	748-755	2
	IVTRIVELL	776-784	5
	LLQYWSQEL	798-806	5
	LYKYKVVKI	482-490	11
	LLQLTVWGI	564-572	12
	WNHTTWMEW	622-630	2
	YKYKVVKIE	483=491	8
	LQARILAVE	575-583	4
	LFSYHRLRD	764-772	8
	IVNRVRQGY	703-711	12
	WQKVGKAMY	426-434	4
	IRGKVQKEY	164-172	8
	VQARQLLSG	538-546	4
	MRVKEKYQH	483-491	22
	LLTRDGGNS	450-460	2
	VERYLKDQQ	582-590	2
	IIFKQSSGG	358-366	1
	WGIKQLQAR	570-578	3
	INMWQKVGK	423-431	13
	VELLGRRGW	781-789	4
	LLELDKWAS	659-667	9
	IMIVGGLVG	685-693	11
	LKPCVKLTP	115-123	8
	IVTHSFNCG	370-378	3
	LCLFSYHRL	762-769	3
NEF	WRFDSRLAF	182-190	25
	VEPDKIEEA	147-155	6
	LVPVEPDKI	144-152	3
	VIGWPTVRE	9-17	4
	VRYPLTFGW	132-140	5
	LWIYHTQGY	111-119	2
	YKLVPVEPD	142-150	3
	MRRAEPAAD	19-27	10
	FPVTPQVPL	67-75	2
	WSKSSVIGW	4-12	2
	YPLTFGWCY	134-142	2
	LTFGWCYKL	136-144	2
	MTYKAAVDL	78-86	2
	WCYKLVPVE	140-149	4
	WQNYTPGPG	123-131	5
	MGGKWSKSS	0-8	2
	FGWCYKLVP	138-146	2
	VRERMRRAE	15-23	7
	IHSQRRQDI	100-108	3
	LRPMTYKAA	75-83	4
	IHSQRRDI	100-107	2
	LIHSQRRQD	99-107	1

4) CONCLUSION

HIV is a major human health problem and a need arises for sensitive and specific diagnostic methods. Currently, in practical pathologists are doing immunoassays as the diagnostic measure, that require purified antigen which is possible by using culture of HIV or recombinant technology. To avoid this problem, bioinformatics is a good approach for prediction of epitopes without using wet laboratory practice. Our approach can help to accelerate virology research. These T-cell epitopes can be chemically synthesized and used as antigen for HIV diagnostic and it can be also used for development of peptide based vaccine.

REFERENCES

1. Weiss, R.A. 1993. How does HIV cause AIDS? Science, 260: (5112); 1273- 79.
2. CDC. Monitoring selected national HIV prevention and care objectives by using HIV surveillance data - United States and 6 U.S. dependent areas - 2010. HIV

Surveillance Supplemental Report 2012; 17 (No. 3, part A).
3. http://nacoonline.org/upload/REPORTS/NACO%20Annual%20Report%202010-11.pdf).
4. Jha, P., Kumar, R., Khera, A., Bhattacharya, M., Arora, P., Gajalakshmi, V., Bhatia, P., Kam, D., Bassani, D. G., Sullivan, A., Suraweera, W., McLaughlin, C., Dhingra,N., Nagelkerke, N. 2010. HIV mortality and infection in India: estimates from nationally representative mortality survey of 1.1 million homes. BMJ, 340: c621.
5. Gilbert, P.B., McKeague, I.W., Eisen, G., Mullins, C., Guéye-Ndiaye, A., Mboup, S., Kanki, P.J. 2003. Comparison of HIV-1 and HIV-2 infectivity from a prospective cohort study in Senegal. Statistics in Medicine, 22 (4): 573-593.
6. Reeves, J.D., Doms, R.W. 2002. Human Immunodeficiency Virus Type 2". Journal of General Virology, 83(6): 1253–65.
7. Petrovsky, N., Silva, D., Brusic, V. 2003. The future for computational modelling and prediction systems in clinical immunology. Novartis Foundation symposium, 254: 23-24.
8. Singh, H., Raghava, G.P.S. 2001. ProPred: prediction of HLA-DR binding sites. Bioinformatics, 17: 1236-37.
9. Singh, H., Raghava, G.P.S. 2003. ProPred1: prediction of promiscuous MHC class-I binding sites. Bioinformatics, 19: 1009-1014.
10. Chaudhary, D.K., Mani I., Singh V. 2012. Identification of T-cell antigenic determinants in the proteins of Human Papillomavirus. In Adv Biotech & Bioinfo, 1(1):79–86.
11. Chaudhary, D.K., Mani, I., Singh V. 2012. Identification of T-cell epitopes in structural proteins of Tick borne encephalitis virus for vaccine development. In App Bio & Pharma Tech, 3(2): 39-43.
12. Somvanshi, P., Singh, V., Seth, P.K. 2008. Prediction of epitopes in hemagglutinin and neuraminidase proteins of Influenza A virus H5N1 strain: A clue for diagnostic and vaccine development OMICS: J Integrative Biol, 12(1): 61–69.
13. Somvanshi, P., Singh, V., Seth, P.K. 2008. In silico prediction of epitopes in virulence proteins of Mycobacterium tuberculosis H37Rv for diagnostic and subunit vaccine design. J Proteomics & Bioinfo, 1(3): 143-153.